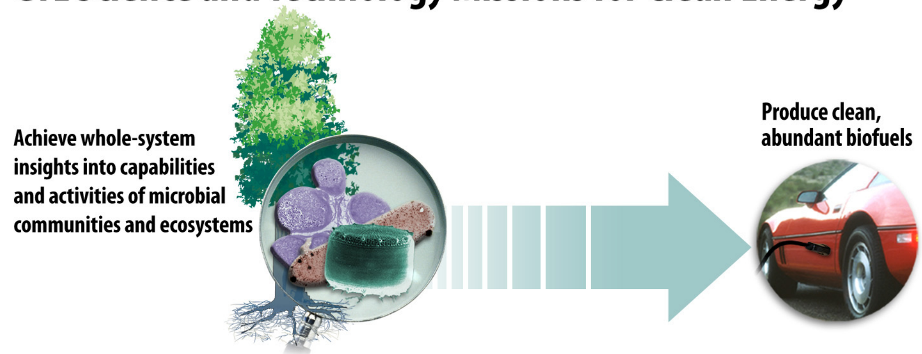
Making Bioethanol Cost Competitive

DOE GENOMICS:GTL
SYSTEMS BIOLOGY FOR ENERGY AND ENVIRONMENT
OFFICE OF SCIENCE
U.S. DEPARTMENT OF ENERGY

doegenomestolife.org

Why Biofuels?

GTL Science and Technology Missions for Clean Energy



Global energy demand is projected to rise rapidly in this century due to population growth, increasing standards of living, and the energy intensity of developing economies. Biologically derived fuels (biofuels) can be renewable, sustainable, and expandable to meet the growing demand. They have the potential to be domestically and globally available for energy security, with most being carbon neutral (introducing no additional carbon to the global carbon cycle) or potentially carbon negative (if coupled with carbon sequestration) and supportable within the current agricultural infrastructure.

Genomics:GTL R&D and Bioethanol

Ethanol produced from cornstarch already is used as a substitute or octane booster for gasoline. Ethanol from this source, however, has limited potential for substantial improvements in volume and cost. Biotechnology offers the promise of dramatically increasing ethanol production using cellulose from such plant waste as post-harvest corn plants and timber residues or from such high-biomass "energy" crops as poplar trees and switchgrasses (see Table 1, Cellulosic Ethanol Goals and Impacts, p. 2). Using cellulose—the most abundant biological material on earth—and other constituents of plant cell walls as feedstocks will allow the development of ethanol as a viable fuel alternative and reduce net CO_2 emissions from the transportation sector. Research in the Genomics:GTL program (GTL) of the Department of Energy (DOE) will lead to increased understanding of the bioprocesses important for cellulosic ethanol production, which will result in increased production efficiencies and reduced costs. GTL's R&D in this area will help bioethanol become a cost-competitive alternative to gasoline in the coming decades (see graph, Projected Ethanol Wholesale Costs, p. 2).

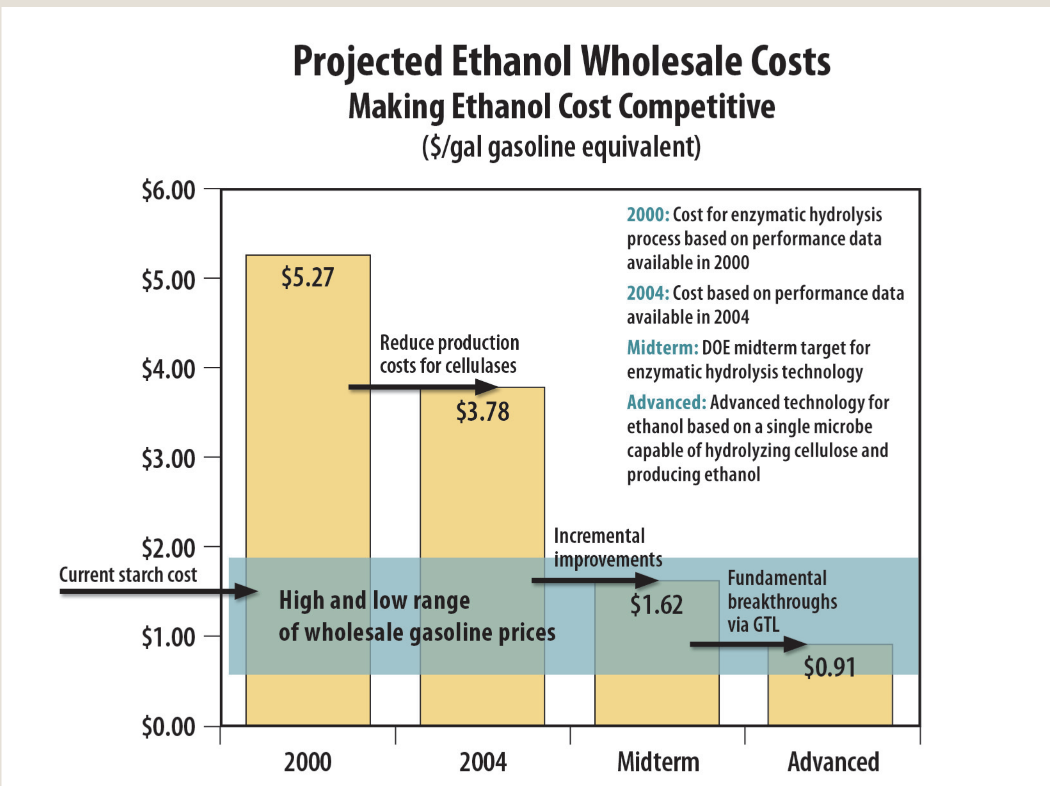
Targets for GTL Research: Improved Enzymes, Integrated Processes

The current commercial approach for converting cellulose into ethanol uses a complicated and expensive multistep process that combines thermochemical and biological methods in large, centralized processing plants. Biomass conversion involves three basic steps: (1) Thermochemical treatment of raw biomass to render the complex polymers (cellulose, hemicellulose, and lignin) more accessible to enzymatic breakdown; (2) production and application of special enzyme preparations (cellulases and hemicellulases) that break down (hydrolyze) plant cell-wall

Table 1. Cellulosic Ethanol Goals and Impacts

Factors	Today–Starch	Interim–Starch and Cellulose	Long Term–Cellulose*
Billion gallons Fossil fuel displaced** CO_2 reduced	4 2% 1.8%	20 10% 9%	30 to 200+ 15 to 100%*** 14 to 90%
Feedstock****	Starch (14% energy yield)	Waste cellulose	Cellulosic energy crops (>37% energy yield)
Process	Starch fermentation Little cellulose processing	Acid decrystallization: Transition to enzymes Cellulases Single sugar metabolism Multiple microbes Some energy crops	Enzyme decrystallization and depolymerization Cellulase and other glycosyl hydrolases Sugar transporters High-temperature functioning Multisugar metabolism Integrated processing Designer cellulosic energy crops Carbon sequestration through plant partitioning
Deployment	Large, central processing	Large, central processing	Distributed or centralized, efficient processing plants
Other impacts: Energy dollars spent at home, third crop for agriculture, land revitalization and stabilization, habitat, soil carbon sequestration, yield per acre roughly tripled (cellulose over corn starch).			

*Enabled by GTL.
**Current U.S. consumption of gasoline is about 137 billion gallons per year, which corresponds to some 200 billion gallons of ethanol (N. Greene et al., *Growing Energy: How Biofuels Can Help End America's Oil Dependence*, Natural Resources Defense Council, New York, 2004). A gallon of ethanol has 2/3 the energy content of a gallon of gasoline.
***Assumes improvements in feedstock yield and supply, efficiency gains in conversion processes and vehicle technologies, and development of fuel infrastructure.
****Adapted from S. J. Smith et al. 2004. *Near-Term US Biomass Potential: Economics, Land-Use, and Research Opportunities*, PNWD-3285, Battelle Memorial Institute, Joint Global Change Research Institute, Baltimore, MD.



Adapted from a 2005 NREL diagram.

constituents (polysaccharides) into a mixture of simple sugars; and (3) fermentation, mediated by bacteria or yeast, to convert these sugars to ethanol. The hydrolysis of cellulose for fermentation is difficult because of its strong, rigid nature. Making the conversion of cellulose to ethanol more economical and practical will require the development of a science base for molecular redesign of numerous enzymes, biochemical pathways, and full cellular systems.

Screening, Analyzing, and Improving Cellulose-Degrading Enzymes

Many microbes and fungi are sources of enzymes called cellulases that can degrade cellulose into its simpler components for processing. GTL research will produce and characterize complex cellulase structures and their functions by analyzing naturally occurring and modified protein and molecular machine variants of essential pathways, understanding the synergistic activity of multiple cellulases, and resolving temperature-sensitivity issues that prevent optimal functioning of cellulase enzymes at fermentative temperatures. GTL will screen a large number of organisms and natural communities to increase the number of enzymes that can be examined. With improved cellulases, we can replace expensive thermochemical processes (see illustration and Table 2 below). Other GTL challenges include characterizing the structures and functions of membrane-bound molecular machines that deliver sugars to the metabolic pathways of fermentative organisms, understanding inefficiencies in conversion of different sugars to ethanol, maintaining large-scale mixed cultures, and improving disease resistance.

Combining Multistep Processes

Understanding cell regulatory processes will be central to incorporating multiple functionalities into a single organism or microbial consortium and enabling optimized overexpression in many related processes. The ultimate innovation, an integrated system combining all key steps in one process using either a single microbe or stable mixed culture, would enable smaller-scale and more cost-effective and energy-efficient distributed processing plants (see chart, p. 4).

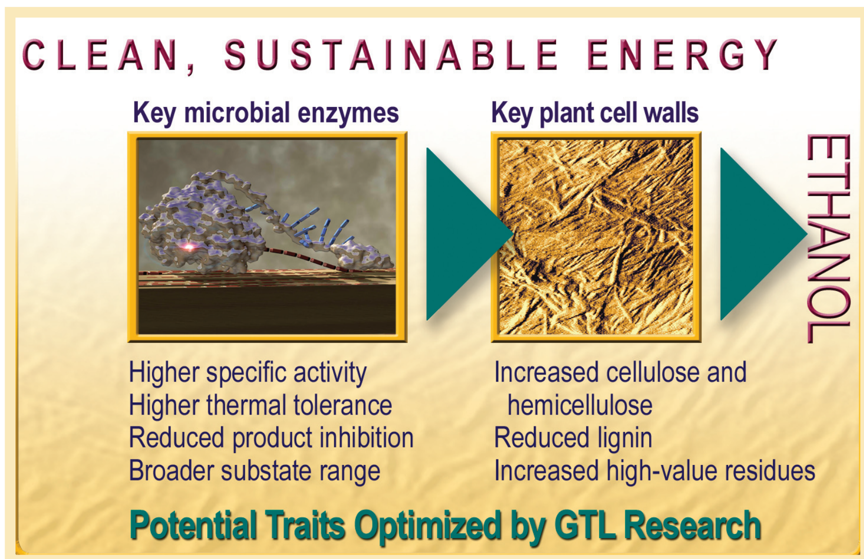
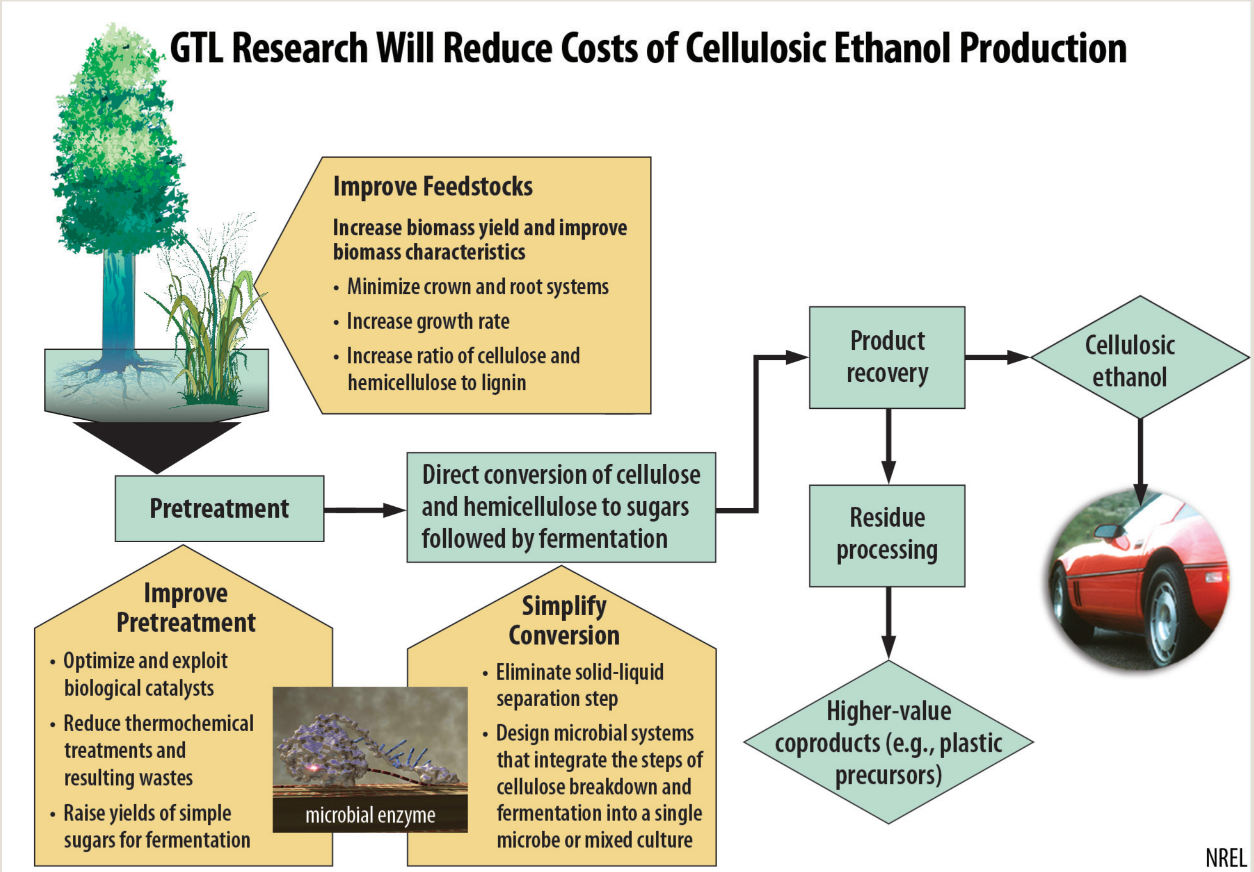


Table 2. Cellulosic Ethanol Challenges, Scale, and Complexity

Research and Analytical Challenges	Scale and Complexity
• Screening of databases for natural variants of cellulases (generally glycosyl hydrolases) and other enzymes or molecular machines in metabolic networks; and characterization of variants • Analysis of modified variants to establish design principles and functional optimization • Modeling and simulation of cellulase, sugar transport, and multiple sugar-fermentation processes and their regulation • Integration of processing steps into single microbes or stable cultures	• Thousands of variants of all enzymes; screening of millions of genes, thousands of unique species and functions • Production and functional analysis of potentially thousands of modified enzymes, hundreds of regulatory processes and interactions • Models at the molecular, cellular, and community levels incorporating signaling, sensing, regulation metabolism, transport, biofilm, and other phenomenology and using massive databases in GTL's knowledgebase • Incorporation of complete cellulose-degradation and sugar-fermentation processes into microbes or consortia—hundreds of metabolic, regulatory, and other interconnected pathways



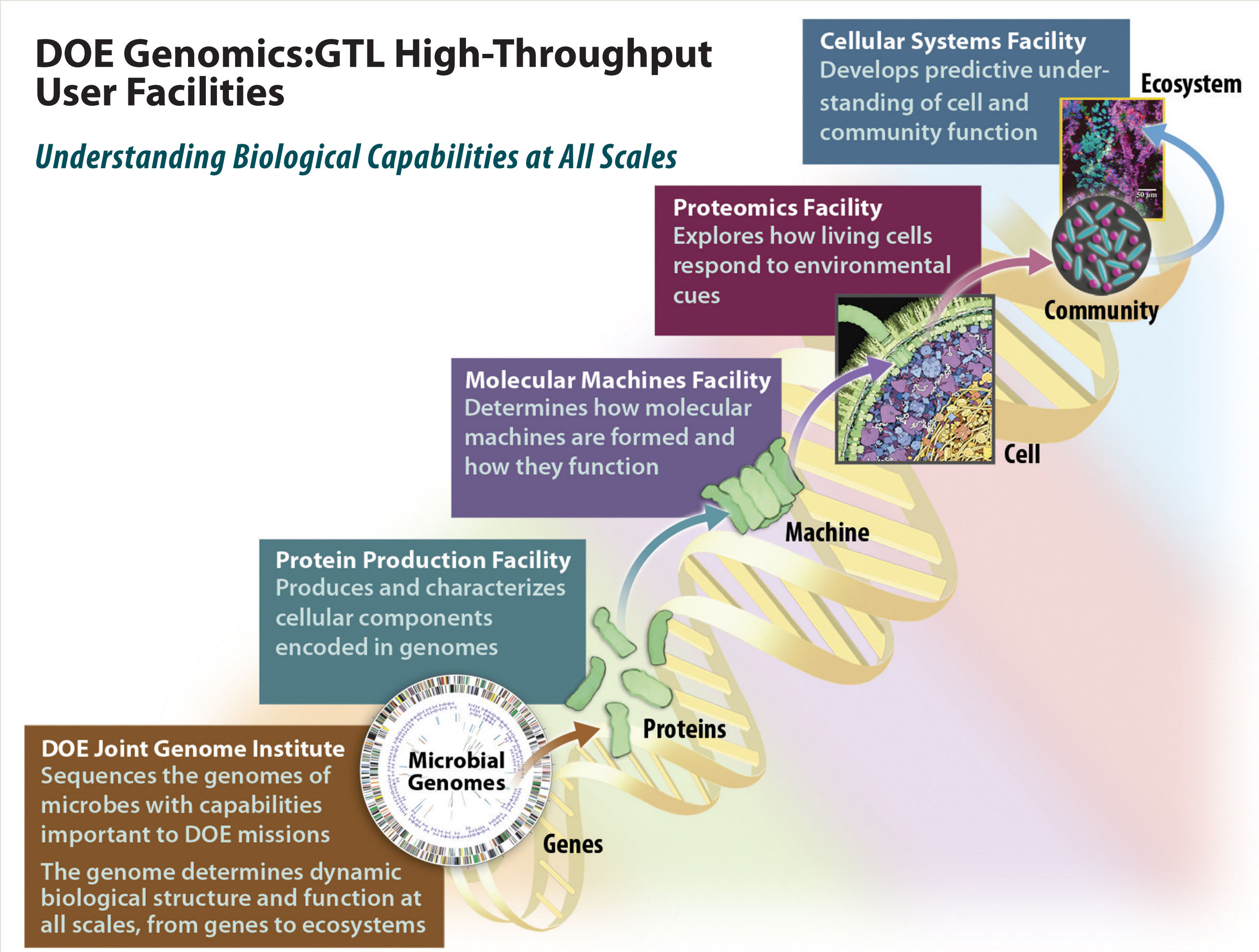
A more complete understanding of enzymes and microbes involved in biomass conversion to ethanol is needed to overcome current inefficiencies in the production process. GTL's bioethanol research targets (shown in the yellow boxes) include improving feedstocks and pretreatment and simplifying conversion, thus enabling development of integrated bioprocessing.

Clostridium thermocellum is a bacterium capable of both hydrolyzing cellulose and fermenting sugars to ethanol, but its yields are poor and conversion is slow. Achieving greater efficiencies with *C. thermocellum* or other microbes will require new methods for modifying them. GTL is establishing a new research paradigm to focus on understanding how microbial systems function and how their interacting pathways influence one another, rather than on studying only a few genes or enzymes.

Planned GTL User Facilities to Enable Timely Applications

Improving understanding of bioethanol production will require a variety of new capabilities including techniques for surveying enzyme diversity; visualizing enzyme systems and cell-wall structures; efficiently producing enzyme systems and membrane proteins; cultivating microbial consortia; integrating transcriptomics, proteomics, and metabolomics; and genetically engineering microorganisms for integrated bioprocessing. Optimization of enzymes for conversion efficiency will require a more detailed understanding of their regulation and activity as a tightly controlled, highly organized system.

The technical challenges of these combined analyses and the scale of systems that must be understood exceed any existing capabilities. New user facilities proposed by GTL would dramatically improve research performance,



throughput, quality, and cost. The graphic, p. 5, describes facilities focused on building an integrated body of knowledge about microbes, from genome interactions through ecosystem changes. Simultaneously studying multiple microbial systems related to DOE missions is powerfully synergistic because enduring biological themes (e.g., similar genes and biochemical pathways) are shared and general principles governing response, structure, and function apply throughout the systems.

High-Value By-Products Further Reduce Costs of Producing Bioethanol

The knowledge gained in GTL will provide the foundation for countless new commercial bioproducts and bioprocesses. A strategy for introducing these technologies to the marketplace could be development of integrated biorefineries capable of producing a suite of products as substitutes for chemical-based fossil feedstocks. Market demand for these high-value alternatives can generate financial returns that make biorefineries commercially competitive, providing a viable base for such lower-value products as transportation fuels. After meeting market opportunities for high-value products, other lower-value products now derived from fossil fuels would be produced, ultimately leading to mass marketing of fuels from biomass resources.

GTL Research Progress

Genome Sequencing — A Prelude to Systems Reengineering

Research to Improve Feedstock Characteristics and Increase Yield

- **DOE's Office of Biological and Environmental Research (BER), as part of an international consortium, funded the sequencing and analysis of the *Populus* tree genome at its Joint Genome Institute (JGI).**
- **BER has sequenced and analyzed the genomes of more than 200 microbes, some of which are relevant to cellulosic ethanol production.**

Research to Improve Efficiency of Cellulose Degradation

- **JGI has determined the genome sequence of white-rot fungus (a lignin degrader), allowing deeper insights into mechanisms of lignin and cellulose degradation. Sequences are the foundation for engineering biological systems or their components to optimize processes for industrial applications.**
- **JGI is determining the sequence of termite hindgut microbial consortia that naturally break down cellulose.**

Since the program began 3 years ago, GTL has supported projects at more than 30 institutions—universities, national labs, and industry. GTL-funded scientists have made significant progress toward creating the next generation of fundamental research, comprising the necessary methodologies and infrastructure. This can now be applied to improve both the yield of feedstock cellulose and the processes for converting cellulose to ethanol.

GTL scientists are tackling the fundamental science required for reengineering biological systems by

- Identifying and characterizing proteins produced by targeted organisms and understanding biochemistries carried out by interacting proteins and other molecules—the molecular machines of life;
- Identifying and understanding how complex gene regulatory networks in cells and communities of cells orchestrate the timing and expression of proteins responsible for growth and other processes;
- Establishing GTL's systems biology computing environment, including modeling and simulating behaviors of biological systems under various conditions; and
- Carrying out pilot projects for eventual research scaleup into facilities to be used by those doing both basic and applied research in DOE mission areas.

Engaging the Bioenergy Research and Industrial Communities

In March 2004, BER organized a workshop to define for the GTL roadmap the science needed to remove biological roadblocks preventing cellulosic ethanol production on a scale sufficient to displace imported fossil fuels. The 2005 roadmap has introductory material, a more detailed appendix, and research scenarios on energy to guide scientists and program managers in optimizing the GTL program around this important mission (more details on p. 7).

In June 2005, a joint workshop was convened with BP America Inc. as coorganizer to define specific research needed to improve biomass feedstock yields. BER and DOE's Office of Energy Efficiency and Renewable Energy are working together on a comprehensive research strategy to improve feedstock yields and lower conversion costs.

Holding several symposia at major meetings, BER has begun a dialog with the Biotechnology Industry Organization to get input on industrial bioenergy needs and facilitate partnering of research scientists with biofuel developers.

Genomics:GTL Roadmap—Setting a Course for Energy Security

The 2005 GTL roadmap (doegenomestolife.org/roadmap/) builds on and expands the first GTL roadmap published in 2001. It traces connections among technical DOE mission objectives and science needs and GTL goals and milestones in biological research, technology, and infrastructure including four world-class user facilities. The roadmap is the result of 3 years of collaboration among hundreds of scientists and technologists via a number of workshops and other activities covering all relevant aspects of science, DOE missions, technologies, and computing. Presenting a baseline for the research facilities, the roadmap is meant to begin the dialogue that will determine their ultimate functionality and form. A vigorous process to refine these ideas and incorporate progress and revolutions as they occur will be central to implementating this plan.

Drawing heavily on the output and insights from some of the workshops, the GTL program has identified gaps in our understanding of microbial processes essential to bioethanol production. Innovation in biotechnology requires basic research that explores a greater variety of enzymes and microorganisms, analyzes enzymes as systems, and determines how certain factors influence biomass degradation or ethanol production. Ongoing attempts to develop and commercialize these technologies have generated several fundamental scientific questions in need of further investigation. They include:

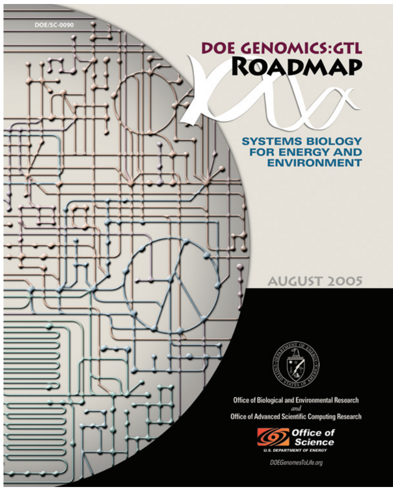
- What is the extent of natural diversity among biomass-degrading and ethanologenic organisms?
- How do soluble enzymes act on an insoluble crystalline substrate?
- How do biomass-degrading enzymes work together as a synergistic system?
- Why are ethanologenic organisms less efficient at using certain sugar substrates?
- How effective are sugar transporters at translocating various sugars across the cell membrane?
- Why do different enzymatic and microbial processes operate optimally at different temperatures?
- What are the requirements for producing and maintaining stable mixed cultures?
- How can we improve tools for genetically engineering microorganisms involved in bioethanol production?

Roadmap Executive Summary

Providing solutions to major national problems, biology and industrial biotechnology will serve as an engine for economic competitiveness in the 21st Century. DOE missions in energy security, environment, and climate are grand challenges for a new generation of biological research. As a mission agency, DOE can bring together resources and expertise in the biological, computing, and physical sciences for the focused and large-scale research effort needed—from scientific investigations to commercialization in the marketplace.

Our investment in genomics over the past 20 years now allows us to rapidly determine and interpret any organism's complete DNA sequence. Because genomics reveals the blueprint for life, it is the launching point for an integrated and mechanistic systems understanding of biological function. It is a new link between biological research and the development of biotechnologies. With genomics data as a starting point, GTL will use a systems biology approach to fundamentally transform the way scientists conduct biological investigations and describe living systems.

GTL's goal is simple in concept but complicated in practice—to reveal how the static information in genome sequences drives the intricate and dynamic processes of life. Through predictive models of these life processes and supporting research infrastructure, we seek to harness the capabilities of microbes and complex microbial communities that are the foundation of the biosphere and sustain all life on earth. Gaining reliable use of microbial processes requires understanding the whole living system, not just genomic DNA sequences or collections of proteins or cell by-products. GTL will study critical microbial properties and processes on three system levels—molecular, cellular, and community—each requiring advances in fundamental capabilities and concepts.



Request copies of the *GTL Roadmap* via doegenomestolife.org/roadmap/

Already, discoveries in the microbial world are changing our view of the origins, limits, and capabilities of life. Unique microbial biochemistries amassed over eons in every niche on the planet now offer a deep and virtually limitless resource that can be applied to help enable biobased solutions for national needs. GTL research will reveal processes by which microbes produce energy, including ethanol and hydrogen, and other capabilities that may be used to clean up environmental contaminants and control the cycling of carbon.

Elucidating the design principles of microbial systems in their diverse environments entails analyses of unprecedented scale and complexity. DOE-relevant microbial systems can have millions of genes and thousands of genetic and regulatory processes and community interactions that underlie diversity and adaptability. Achieving GTL goals requires a major advance in our ability to measure the phenomenology of living systems and to incorporate their operating principles into computational models and simulations that accurately represent biological systems—the ultimate level of integrated understanding generated by GTL research.

To make GTL science and biological research more broadly tractable, timely, and affordable, GTL will develop four user facilities, delivering economies of scale and enhanced performance. These facilities will provide the advanced technologies and state-of-the-art computing needed to better understand microbial genomic potential, cellular responses, regulation, and community behaviors in any environment. Making these resources available to the broad research and technology-development communities will democratize access to forefront scientific resources and enlist an expanded community in exciting science for national needs.

Central to the success of the GTL program are computing and information technologies that will help surmount the barrier of complexity now preventing us from deducing biological function directly from genome sequence. GTL will create an integrated computational environment that will link experimental data of unprecedented quantity and dimensionality with theory, modeling, and simulation to uncover fundamental biological principles and to develop and test systems theory for biology.

The GTL Roadmap traces the path from DOE mission science through systems microbiology to the promise of emerging technologies, integrated computing, and a new research infrastructure. It describes opportunities, research strategies, and solutions at the nexus of the challenges of this new science as applied to microbes and the complexities of mission problems.

Office of Science
U.S. DEPARTMENT OF ENERGY

Genomics:GTL is a partnership between the Office of Biological and Environmental Research and the Office of Advanced Scientific Computing Research within the Department of Energy's Office of Science.

John Houghton
Office of Biological and Environmental Research
U.S. Department of Energy
301.903.8288, john.houghton@science.doe.gov

Gary Johnson
Office of Advanced Scientific Computing Research
U.S. Department of Energy
301.903.5800, gary.johnson@science.doe.gov

DOE Office of Science: www.science.doe.gov

October